IN THE CLAIMS:

Status of the claims:

Claims 19, 24, 32 and 33 are previously amended;

Claims 1-18, 20-23, 25-31 and 34-43 are original.

1.   (Original)   A system for generating a description record from multimedia information, comprising:

(a)   at least one multimedia information input interface receiving said multimedia information;

(b)   a computer processor, coupled to said at least one multimedia information input interface, receiving said multimedia information therefrom, processing said multimedia information by performing object extraction processing to generate multimedia object descriptions from said multimedia information, and processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions indicative of an organization of said object descriptions, wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated for content embedded within said multimedia information; and

(c)   a data storage system, operatively coupled to said processor, for storing said at least one description record.

2.   (Original)   The system of claim 1, wherein said multimedia information comprises image information, said multimedia object descriptions comprise image object descriptions, and said multimedia object hierarchy descriptions comprise image object hierarchy descriptions.

3.   (Original)   The system of claim 2, wherein said object extraction processing comprises:
      (a)   image segmentation processing to segment each image in said image information into regions within said image; and
      (b)   feature extraction processing to generate one or more feature descriptions for one or more of said regions;
whereby said generated object descriptions comprise said one or more feature descriptions for one or more of said regions.

4.   (Original)   The system of claim 3, wherein said one or more feature descriptions are selected from the group consisting of text annotations, color, texture, shape, size, and position.

5.   (Original)   The system of claim 2, wherein said object hierarchy processing comprises physical object hierarchy organization to generate physical object hierarchy descriptions of said image object descriptions that are based on spatial characteristics of said objects, such that said image object hierarchy descriptions comprise physical descriptions.

6.   (Original)   The system of claim 5, wherein said object hierarchy processing further comprises logical object hierarchy organization to generate logical object hierarchy descriptions of said image object descriptions that are based on semantic characteristics of said objects, such that said image object hierarchy descriptions comprise both physical and logical descriptions.

7.   (Original)   The system of claim 6, wherein said object extraction processing comprises:
      (a)   image segmentation processing to segment each image in said image information into regions within said image; and

(b)   feature extraction processing to generate object descriptions for one or more of said region;

and wherein said physical hierarchy organization and said logical hierarchy organization. generate hierarchy descriptions of said object descriptions for said one or more of said regions.

8.   (Original)   The system of claim 7, further comprising an encoder receiving said image object hierarchy descriptions and said image object descriptions, and encoding said image object hierarchy descriptions and said image object descriptions into encoded description information, wherein said data storage system is operative to store said encoded description information as said at least one description record.

9.   (Original)   The system of claim 1, wherein said multimedia information comprises video information, said multimedia object descriptions comprise video object descriptions including both event descriptions and object descriptions, and said multimedia hierarchy descriptions comprise video object hierarchy descriptions including both event hierarchy descriptions and object hierarchy descriptions.

10.   (Original)   The system of claim 9, wherein said object extraction processing comprises:

(a)   temporal video segmentation processing to temporally segment said video information into one or more video events or groups of video events and generate event descriptions for said video events,

(b)   video object extraction processing to segment said one or more video events or groups of video events into one or more regions, and to generate object descriptions for said regions; and

(c)   feature extraction processing to generate one or more event feature descriptions for said one or more video events or groups of video events, and one or more object feature descriptions for said one or more regions;

wherein said generated video object descriptions include said event feature descriptions and said object descriptions.

11.   (Original)   The system of claim 10, wherein said one or more event feature descriptions are selected from the group consisting of text annotations, shot transition,

camera motion, time and key frame, and wherein said one or more object feature descriptions are selected from the group consisting of color, texture, shape, size, position, motion, and time.

12.   (Original)   The system of claim 9, wherein said object hierarchy processing comprises physical event hierarchy organization to generate physical event hierarchy descriptions of said video object descriptions that are based on temporal characteristics of said video objects, such that said video hierarchy descriptions comprise temporal descriptions.

13.   (Original)   The system of claim 12, wherein said object hierarchy processing further comprises logical event hierarchy organization to generate logical event hierarchy descriptions of said video object descriptions that are based on semantic characteristics of said video objects, such that said hierarchy descriptions comprise both temporal and logical descriptions.

14.   (Original)   The system of claim 13, wherein said object hierarchy processing further comprises physical and logical object hierarchy extraction processing, receiving said temporal and logical descriptions and generating object hierarchy descriptions for video objects embedded within said video information, such that said video hierarchy descriptions comprise temporal and logical event and object descriptions.

15.   (Original)   The system of claim 14, wherein said object extraction processing comprises:

    (a)   temporal video segmentation processing to temporally segment said video information into one or more video events or groups of video events and generate event descriptions for said video events,

    (b)   video object extraction processing to segment said one or more video events or groups of video events into one or more regions, and to generate object descriptions for said regions; and

    (c)   feature extraction processing to generate one or more event feature descriptions for said one or more video events or groups of video events, and one or more object feature descriptions for said one or more regions;

wherein said generated video object descriptions include said event feature descriptions and said object descriptions, and wherein said physical event hierarchy organization and said logical event hierarchy organization generate hierarchy descriptions from said event feature descriptions, and wherein said physical object hierarchy organization and said logical object hierarchy organization generate hierarchy descriptions from said object feature descriptions

16. (Original) The system of claim 15, further comprising an encoder receiving said video object hierarchy descriptions and said video object descriptions, and encoding said said video object hierarchy descriptions and said video object descriptions into encoded description information, wherein said data storage system is operative to store said encoded description information as said at least one description record.

17. (Original) A method for generating a description record from multimedia information, comprising the steps of:

    (a)   receiving said multimedia information;

    (b)   processing said multimedia information by performing object extraction processing to generate multimedia object descriptions from said multimedia information;

    (c)   processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions indicative of an organization of said object descriptions, wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated for content embedded within said multimedia information; and

    (d)   storing said at least one description record.

18.   (Original)   The method of claim 17, wherein said multimedia information comprises image information, said multimedia object descriptions comprise image object descriptions, and said multimedia object hierarchy descriptions comprise image object hierarchy descriptions.

19.   (Previously amended) The method of claim 18, wherein said object extraction processing step comprises the sub-steps of:

(a)   image segmentation processing to segment each image in said image information into regions within said image; and

(b)   feature extraction processing to generate one or more feature descriptions for one or more of said regions;

whereby said generated image object descriptions comprise said one or more feature descriptions for one or more of said regions.

20.   (Original)   The method of claim 19, wherein said one or more feature descriptions are selected from the group consisting of text annotations, color, texture, shape, size, and position.

21.   (Original)   The method of claim 18, wherein said step of object hierarchy processing includes the sub-step of physical object hierarchy organization to generate physical object hierarchy descriptions of said image object descriptions that are based on spatial characteristics of said objects, such that said image hierarchy descriptions comprise physical descriptions.

22.   (Original)   The method of claim 21, said step of object hierarchy processing further includes the sub-step of logical object hierarchy organization to generate logical object hierarchy descriptions of said image object descriptions that are based on semantic characteristics of said objects, such that said image object hierarchy descriptions comprise both physical and logical descriptions.

23.   (Original)   The method of claim 22, wherein said step of object extraction processing further includes the sub-steps of:

(a)   image segmentation processing to segment each image in said image information into regions within said image; and

(b)   feature extraction processing to generate object descriptions for one or more of said region;

and wherein said physical object hierarchy organization sub-step and said logical object hierarchy organization sub-step generate hierarchy descriptions of said object descriptions for said one or more of said regions.

24.   (Previously amended) The method of claim 18, further comprising the step of encoding said image object descriptions and said image object hierarchy descriptions into encoded description information prior to said data storage step.

25.   (Original)   The method of claim 17, wherein said multimedia information comprises video information, said multimedia object descriptions comprise video object descriptions including both event descriptions and object descriptions, and said multimedia hierarchy descriptions comprise video object hierarchy descriptions including both event hierarchy descriptions and object hierarchy descriptions.

26.   (Original)   The method of claim 25, wherein said step of object extraction processing comprises the sub-steps of:

(a)   temporal video segmentation processing to temporally segment said video information into one or more video events or groups of video events and generate event descriptions for said video events,

(b)   video object extraction processing to segment said one or more video events or groups of video events into one or more regions, and to generate object descriptions for said regions; and

(c)   feature extraction processing to generate one or more event feature descriptions for said one or more video events or groups of video events, and one or more object feature descriptions for said one or more regions;

wherein said generated video object descriptions include said event feature descriptions and said object descriptions.

27.   (Original)   The method of claim 26, wherein said one or more event feature descriptions are selected from the group consisting of text annotations, shot transition, camera motion, time and key frame, and wherein said one or more object feature

descriptions are selected from the group consisting of color, texture, shape, size, position, motion, and time.

28.   (Original)   The method of claim 25, wherein said step of object hierarchy processing includes the sub-step of physical event hierarchy organization to generate physical event hierarchy descriptions of said video object descriptions that are based on temporal characteristics of said video objects, such that said video hierarchy descriptions comprise temporal descriptions.

29.   (Original)   The method of claim 28, wherein said step of object hierarchy processing further includes the sub-step of logical event hierarchy organization to generate logical event hierarchy descriptions of said video object descriptions that are based on semantic characteristics of said video objects, such that said hierarchy descriptions comprise both temporal and logical descriptions.

30.   (Original)   The method of claim 29, wherein said step of object hierarchy processing further comprises the sub-step physical and logical object hierarchy extraction processing, receiving said temporal and logical descriptions and generating object hierarchy descriptions for video objects embedded within said video information, such that said video hierarchy descriptions comprise temporal and logical event and object descriptions..

31.   (Original)   The method of claim 30, wherein said step of object extraction processing comprises the sub-steps of:

    (a)   temporal video segmentation processing to temporally segment said video information into one or more video events or groups of video events and generate event descriptions for said video events,

    (b)   video object extraction processing to segment said one or more video events or groups of video events into one or more regions, and to generate object descriptions for said regions; and

    (c)   feature extraction processing to generate one or more event feature descriptions for said one or more video events or groups of video events, and one or more object feature descriptions for said one or more regions;

wherein said generated video object descriptions include said event feature descriptions and said object descriptions, and wherein said physical event hierarchy organization and said logical event hierarchy organization generate hierarchy descriptions from said event feature descriptions, and wherein said physical object hierarchy organization and said logical object hierarchy organization generate hierarchy descriptions from said object feature descriptions.

32.   (Previously amended) The method of claim 31, further comprising the step of encoding said video object descriptions and said video object hierarchy descriptions into encoded description information prior to said data storage step.

33.   (Previously amended) A computer readable media containing digital information with at least one multimedia description record describing multimedia content for corresponding multimedia information, the description record comprising:

    (a)   one or more multimedia object descriptions, generated by performing

            object extraction processing, said object descriptions describing

            corresponding multimedia objects;

    (b)   one or more features characterizing each of said multimedia object

            descriptions; and

    (c)   one or more multimedia object hierarchy descriptions indicative of an

            organization of said object descriptions, if any, relating at least a portion of

            said one or more multimedia objects in accordance with one or more

            characteristics.

34.   (Original)   The computer readable media of claim 33, wherein said multimedia information comprises image information, said multimedia objects comprise image objects, said multimedia object descriptions comprise image object descriptions, and said multimedia object hierarchy descriptions comprise image object hierarchy descriptions.

35. (Original) The computer readable media of claim 34, wherein said one or more features are selected from the group consisting of text annotations, color, texture, shape, size, and position.

36. (Original) The computer readable media of claim 34, wherein said image object hierarchy descriptions comprise physical object hierarchy descriptions of said image object descriptions based on spatial characteristics of said image objects.

37. (Original) The computer readable media of claim 36, wherein said image object hierarchy descriptions further comprises logical object hierarchy descriptions of said image object descriptions based on semantic characteristics of said image objects.

38. (Original) The computer readable media of claim 33, wherein said multimedia information comprises video information, said multimedia objects comprise events and video objects, said multimedia object descriptions comprise video object descriptions including both event descriptions and object descriptions, said features comprise video event features and video object features, and said multimedia hierarchy descriptions comprise video object hierarchy descriptions including both event hierarchy descriptions and object hierarchy descriptions .

39. (Original) The computer readable media of claim 38, wherein said one or more event feature descriptions are selected from the group consisting of text annotations, shot transition, camera motion, time and key frame, and wherein said one or more object feature descriptions are selected from the group consisting of color, texture, shape, size, position, motion, and time..

40. (Original) The computer readable media of claim 38, wherein said event hierarchy descriptions comprise one or more physical hierarchy descriptions of said events based on temporal characteristics.

41. (Original) The computer readable media of claim 40, wherein said event hierarchy descriptions further comprise one or more logical hierarchy descriptions of said events based on semantic characteristics.

42.   (Original)   The computer readable media of claim 38, wherein said object hierarchy descriptions comprise one or more physical hierarchy descriptions of said objects based on temporal characteristics.

43.   (Original)   The computer readable media of claim 39, wherein said object hierarchy descriptions further comprise one or more logical hierarchy descriptions. of said objects based on semantic characteristics.